Hum Genet (1989) 81:214–220

Human Genetics

Isolation and characterization of the complete human β-myosin heavy chain gene

Klaus W. Diederich, Ingrid Eisele, Thomas Ried, Thomas Jaenicke, Peter Lichter*, and Hans-Peter Vosberg

Max-Planck-Institut für medizinische Forschung, Abteilung für Molekulare Biologie, Jahnstrasse 29, D-6900 Heidelberg, Federal Republic of Germany

Summary. The entire gene coding for the human β-myosin heavy chain has been isolated from genomic EMBL3A phage libraries by chromosomal walking starting from clone gMHC-1, reported earlier (Appelhans and Vosberg 1983). gMHC-1 has been shown to carry coding information for the C-terminal two-thirds of β-myosin heavy chain, which is expressed in cardiac muscle and in slow skeletal muscle fibers (Lichter et al. 1986). Three DNA clones were identified as overlapping with gMHC-1 by restriction mapping and DNA sequencing. They span a 30-kb region in the genome. About 22 kb extend from the initiation codon ATG to the poly(A) addition site. The clones include about 4 kb of 5′ flanking sequences upstream of the promoter. Comparisons of β- and α-myosin heavy chain sequences indicate that gene duplication of the cardiac myosin heavy chain isogenes preceded the mammalian species differentiation.

Introduction

Muscle is a particularly intriguing biological system for studies of gene expression in cells, tissues, and whole organisms. Thus, separate stages can be distinguished in vivo during the development of muscle from early embryonic to adult life (Caplan et al. 1983). Also, different expression programs exist in the adult vertebrate organism in individual muscle tissues (Buckingham 1985; Bandman 1985). And finally, these programs are not fixed, but vary depending on changing internal and external conditions (Blau et al. 1985; Whalen et al. 1985; Izumo et al. 1986). The plasticity of muscle is accompanied by a considerable isoform variability of muscle-specific proteins. The mechanisms responsible for the various proteins are based on multigene families of nonallelic, but related, genes (Buckingham 1985; Emerson and Bernstein 1987), on alternative splicing of primary muscle gene transcripts (Breitbart et al. 1987), and possibly also on posttranslational protein modifications (Bandman et al. 1982).

A major component of the contractile apparatus is the myosin molecules, which consist in their monomeric form of two heavy chains (molecular weight about 200000) and four light chains (molecular weights between 16000 and 27000). The native units assemble spontaneously to produce the multimeric thick filaments of muscle. Myosin heavy chains (HCs) have been identified in multiple unique isoforms in embryonic, neonatal, and adult muscle (Weydert et al. 1983; Whalen et al. 1985). In addition, different adult muscle tissues (fast, slow, cardiac, and others) contain distinct versions of the myosin heavy chains. Their expression is controlled by hormones (Butler-Browne and Whalen 1984; Whalen et al. 1985; Everett et al. 1986), by innervation (Pette and Vrbova 1985; Ecob-Prince et al. 1986), and by the physical state of the muscle (McDermott et al. 1985; Hoffman et al. 1986; Buttrick et al. 1986).

The different isoforms of sarcomeric myosin HCs are coded by a multigene family (Nguyen et al. 1982; Leinwand et al. 1983a) that in man and mouse is split; in man, one locus for cardiac myosin HC genes is on chromosome 14 (Saez and Leinwand 1986; Hiller 1986) and a second locus for the sekeltal myosin HC genes, on chromosome 17 (Leinwand et al. 1983b; Rappold and Vosberg 1983; Edwards et al. 1985). It is not known how many different sarcomeric myosin HC genes exist. A minimum estimate for mammalian organisms is eight (Buckingham 1985).

We previously described genomic DNA clone gMHC-1 (Appelhans and Vosberg 1983), which was subsequently shown to comprise part of a myosin HC gene expressed in heart and skeletal muscle (Lichter et al. 1986). On the basis of partial DNA sequence data we were able to demonstrate that this clone contains about 65% of the β-myosin HC gene, including the 3′ end. To complete the entire gene, we isolated the missing parts by chromosomal walking. Four overlapping clones cover the entire gene together with 5′ and 3′ flanking sequences spanning a 30-kb region. These clones make it possible to study the regulation of the β-myosin HC gene in muscle tissues, in myoblast differentiation in vitro, and in diseased muscle.

Methods

Preparation of DNA from leukocytes

Twenty milliliters of venous blood and 20 mg EDTA were mixed with 60 ml buffer containing 155 m*M* NH_4Cl, 10 m*M* $KHCO_3$, and 0.1 m*M* EDTA (pH 7.3–7.4) and cooled in ice. Hemolysis occurred within 30 min. The intact leukocytes were pelleted for 10 min in an SS34 rotor at 1500 rpm (4°C). Cells were washed with 10 ml buffer containing 75 m*M* NaCl and 25 m*M* EDTA, pH 7.8. After resuspension of the cells in 5 ml

* *Present address:* Yale University School of Medicine, Department of Human Genetics, New Haven, CT 06510, USA

Offprint requests to: H.-P. Vosberg

of the same buffer, proteinase K was added to a concentration of 100 µg/ml and SDS to 1%. After 2 h at 37°C, DNA was extracted with phenol and finally precipitated with isopropanol. We obtained about 500 µg DNA from 20 ml blood. (The hemolysis protocol is from Dr. Dareshwar, Bombay).

Construction of genomic phage libraries

Genomic *Mbo*I fragments, obtained by limited digestion, were treated with alkaline phosphatase and ligated with *Bam*HI-cleaved EMBL3A vector arms (Frischauf et al. 1983). Ligation products were packaged using extracts of Genofit (Geneva) or Boehringer Mannheim. Phages were plated on *E. coli* strain NM 539. This strain carries the prophage P2 and affords the powerful *spi* selection system: Only phages containing foreign inserts are propagated (Karn et al. 1980). Two libraries were constructed separately by ligating (library 1) 200 ng restricted genomic DNA to 300 ng vector DNA and (library 2) 4.5 µg genomic DNA to 6 µg vector DNA. The total plaque yields for the two libraries were about 5×10^5 plaque forming units (pfu) and 1.5×10^6 pfu, respectively. The probabilities to find a 15-kb single-copy insert within these plaque collections were 90% and 99% (Clarke and Carbon 1976). Recombinant phages were harvested directly from the plates.

Phage growth conditions and plaque screening

λ-Phages were grown either in suspension according to Silhavy et al. (1984) or as lysates harvested from plates as described earlier (Appelhans and Vosberg 1983). The plaque-screening protocol was essentially that of Benton and Davis (1977). We used 13.5-cm plates with approximately 2×10^4 pfu per plate. Altogether 20 plates per library were screened. Radioactive labeling of probes [$\alpha(^{32}P)$dCTP with 111 TBq/mmol (1 TBq = 10^{12} decays/s)] was achieved using the random primer extension protocol of Feinberg and Vogelstein (1983). Filters were washed in 0.1 × SSC, 0.1% SDS for 2 h at 68°C.

Restriction mapping and sequencing

Restriction maps of phage DNA or fragments thereof cloned in pBR325 or pUC19 were established using standard techniques (Maniatis et al. 1982). Sequencing of the 5′ terminal fragment of the genomic clone gMHC-1 (designated EcoG) was according to Maxam and Gilbert (1980). All other sequences were obtained by the dideoxy chain termination method (Sanger et al. 1977) using (^{35}S)dATP. In addition to the general sequencing primers of the pUC system we used β-myosin HC-specific oligonucleotide primers, which were synthesized according to the phosphoamidite triester method (Caruthers et al. 1987).

Sequence analysis

Alignment of the first translated exon of the β-myosin HC gene was done using the program HD-MAXHOM by Dr. C. Sander, Heidelberg. Open reading frames and splice sites were identified by a translation program written by K.W.D.

Results

Three EMBL 3A phage clones with myosin heavy chain gene sequences were isolated from two independently constructed genomic phage libraries, libraries 1 and 2. For clone selection two probes were used, both containing single-copy sequences of the previously characterized clone gMHC-1 (Appelhans and Vosberg 1983). One of these probes was the 5′ terminal *Eco*RI fragment of this clone (652 bp, designated EcoG), and the other included parts of the last exon with the entire 3′ non-translated region of the β-myosin HC gene (a 412-bp *Sma*I fragment, designated EcoFIII). No physical size selection of DNA fragments was involved in the construction of the libraries. Instead, digestion conditions were controlled to yield an average fragment length of about 15–20 kb.

From library 1 we isolated two clones that hybridized to both probes (these clones were not investigated further) and one that hybridized to EcoG only. Because this clone did not hybridize with sequences 1 kb downstream of EcoG, it was assumed that it extends approximately from the middle of the gene over a distance of about 14 kb to its 5′ end. This clone was designated gMHC-2. Library 2 contained one clone that was recognized by the EcoG probe, but not by EcoFIII, and a second clone, which was labelled by EcoFIII, but only weakly by EcoG. These two clones were designated gMHC-3 and gMHC-4, respectively. The insert lengths of these EMBL3A clones were about 14 kb. Restriction with the enzymes *Eco*RI, *Bam*HI, and *Hin*dIII produced the following results (see Fig. 1): With gMHC-2 and -3, two common *Eco*RI fragments (3.2 kb and 1.4 kb) and one common *Bam*HI fragment

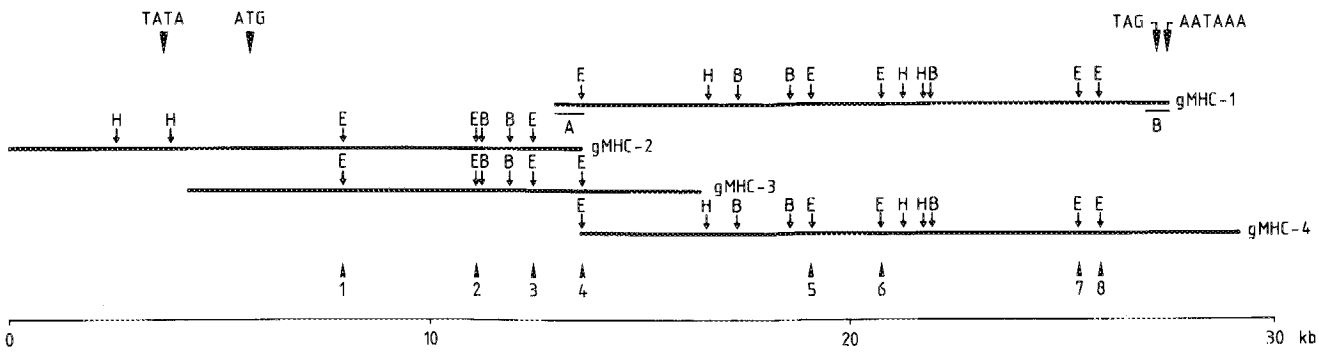
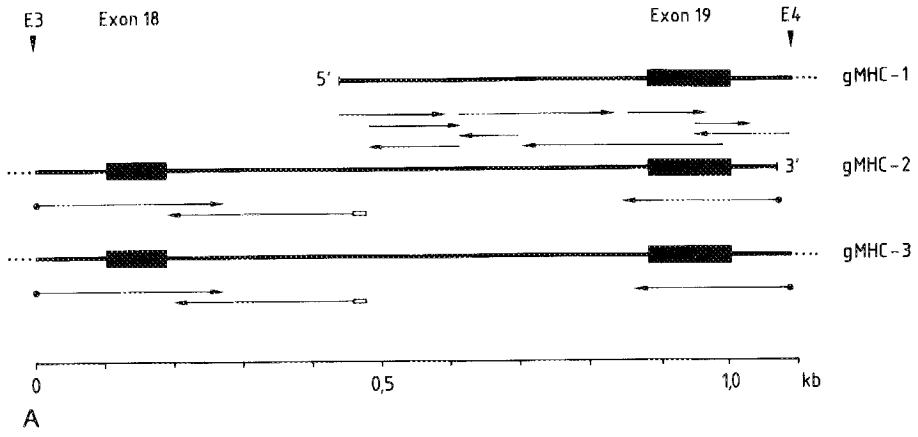


Fig. 1. Restriction maps of the genomic clones gMHC-1 to -4. The restriction sites for *Eco*RI (the eight sites are numbered 1 to 8), *Hin*dIII *(H)*, and *Bam*HI *(B)* are shown. Positions of the promoter *(TATA)*, the translation start codon *(ATG)*, the stop codon *(TAG)*, and the polyadenylation signal *(AATAAA)* are marked by *arrows*. Fragments of gMHC-1 that served as screening probes A (EcoG) and B (EcoFIII) are indicated as *thin lines*. *Lengths* are indicated in kb




```
GAATTCTCTCCATCTCTCTTTTCCTTCCTTCTTCTCCTCTCTTCTTCCTGCATCTCTTTCTGGCATTTTCTTTAC
TTTTCTCTATTGCGATTTTTGGCCACAGGAAAATCTGAACAAGCTGATGACCAACTTGCGCTCCACCCATCCCCA  150
                    E  N  L  N  K  L  H  T  N  L  R  S  T  H  P  H
                                                                    Exon 18
CTTTGTACGTTGTATCATCCCTAATGAGACAAAGTCTCCAGGTGAGGCCACAAACTAGGGCCAACCCACTGCTGG
 F  V  R  C  I  I  P  N  E  T  K  S  P

GCATACACCGCCCTGGGAACAGGACCTCCTAGGACATCTCCCTACCTACCACCACAGTGGCTTTCCAAACCACAT  300
TCTTCTCCTCCCATCCCGCATGAGGCCTCCCAGGGTACCTCCCTCTCAGCGCTACCTTGCACTTGCAGTGTGTAA
TGAGGGTGATGTGTAAGGCACTCTGATGGACATTACCTCATCAGTGGCCCTACAACCCCCTGGATAGCTGGTTGT  450
                                                    └→ gMHC 1
CCAAGACAGAATCTTAGTTCAACTTTAAATTTACCTGCAGAACCACTTAATCCCCCAATGAGACATATTCAAATT
GTTAATCCATTTTTGGCTAATTAGGAGACAGAGGCTCAGGGACCTGGGTTCAAGACCTGGCTCTGTTCTTGATCA  600
GCTTAGTGAACTTGGCCTGTCCCGTAACCCCTCTGGGCTTAATTTACTCATCTAGAACATTGGATATATCTGTTC
TGGCTACCACCTATGATAATAGGTGTGAATCATATAAAAAAGGAATGTGAAAGCCTTTTAGTTAAAAATGTGGCT  750
AGAAGAAAATGAAACAAATGATTATTACTAACTGACTTGCTAAGATTACAAGCTAATCAGTGACAAAGCCAGGAT
CAGAACCCAGAACTTCAGTCCAGTGTTCTCACAGACTCCTCCTACTTCCTTCTTGCCACAGGGGTGATGGACAAC  900
                                                    G  V  M  D  N
                                                                    Exon 19
CCCCTGGTCATGCACCAGCTGCGCTGCAATGGTGTGCTGGAGGGCATCCGCATCTGCAGGAAAGGCTTCCCCAAC
 P  L  V  M  H  Q  L  R  C  N  G  V  L  E  G  I  R  I  C  R  K  G  F  P  N

CGCATCCTCTACGGGGACTTCCGGCAGAGGTGGGTATGAGGGTGCACCAGAGCTCATAGAACAGGGGGAGCCAGG 1050
 R  I  L  Y  G  D  F  R  Q  R

CTGCCCTGATGGGAATGGGATCTGCAGGTGACCCTAGAATTC
                 gMHC 2 ←┘
B
```

Fig. 2A, B. The overlapping region of clones gMHC-1, -2, and -3. This region is flanked by the *Eco*RI sites 3 and 4 (*E3* and *E4* in Fig. 1). **A** Sequencing strategy. The *arrows* underneath gMHC-1 indicate sequences determined according to Maxam and Gilbert (1980). gMHC-2 and -3 sequences were obtained according to Sanger et al. (1977) with pUC primers *(dots)* or synthetic β-myosin HC-specific primers *(open boxes)*. Exon 18 and exon 19, shown as *dark boxes*, are numbered according to the rat embryonic myosin HC gene (Strehler et al. 1986). Fragment lengths in kb. **B** Sequence of λ-GMHC, 1, 2 and 3 in the common *Eco*RI fragment. DNA and protein sequences. *Underlined* are the *Alu*I and *Mbo*I restriction sites. The two sites delineate the 5′ end of gMHC-1 and the 3′ end of gMHC-2 respectively. The amino acid sequences coded by exons 18 and 19 are indicated

(0.7 kb) were obtained. Digestion of the clones gMHC-1 and gMHC-4 produced common fragments with *Eco*RI (5.4 kb, 4.6 kb, 1.6 kb, and 0.5 kb), *Hin*dIII (4.6 kb and 0.5 kb), and *Bam*HI (3.3 kb and 1.2 kb). The order of the restriction sites in the respective clones is summarized in Fig. 1. These results indicate that gMHC-2 and -3, and gMHC-1 and -4 are pairs of overlapping genomic clones.

That the two pairs are mutually overlapping and thus represent a 30-kb contiguous genomic segment was shown by sequence analysis of the region between the *Eco*RI sites 3 and 4 (1092 bp; indicated in Fig. 1). The sequencing strategy is shown in Fig. 2a. Only gMHC-3 contains this region in its entire length. Clone gMHC-1 lacks 440 bp from the 5′ end and gMHC-2 lacks 20 bp from the 3′ end. Six hundred seventy base pairs common to all three clones were found to be identical by complete sequencing of the respective gMHC-1 region and by partial sequencing of about 220 bp at the 3′ ends of the respective regions of gMHC-2 and -3. (Data for gMHC-1 are taken in part from Lichter et al. 1986 and Lichter 1986.) The 440 bp not present in gMHC-1 were found to be identical in both clones, gMHC-2 and -3 (Fig. 2b). gMHC-2 terminates at its 3′ end with an *Mbo*I cleavage site. This site is present in gMHC-1 and -3 (underlined in Fig. 2b), which both extend beyond this restriction sequence. An *Alu*I cleavage site at the 5′ end of gMHC-1 is present in gMHC-2 and -3 (underlined in Fig. 2b). The regions between the *Eco*RI sites 3 and 4 include two β-myosin HC exons, exons 18 and 19. [The exon numbering is according to that for the rat embryonic myosin HC gene (Strehler et al. 1986).] Exon 19 codes for the functionally important "active thiol" region of the myosin head. Clone gMHC-4 was not sequenced. However, the restriction mapping data indicate firmly that this clone overlaps extensively with gMHC-1. It terminates at its 5′ end just 20 bp 5′ to the *Eco*RI site 4.

```
                    putative splice site
                    ↓       Nco I
hum ß CTCACTCCAGGCACAGCCATGGGAGATTCGGAGATGGCAGTCTTTGGGGCTGCCGCCCCCTACCTGCGCAAGTCA
hum α -C--GGGA-*----CAAG---ACC---AGCC-------T-A------**-A*-G----AG-----C---------

hum ß                    M  G  D  S  E  M  A  V  F  G  A  A  A  P  Y  L  R  K  S
rab ß                    -  -  -  A  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -
hum α                    -  T  -  -  Q  -  -  D  -  -  *  T  -  Q  -  -  -  -  -
rab α                    -  T  -  A  Q  -  -  D  -  -  -  -  -  -  -  -  -  -  -

hum ß GAGAAGGAGCGGCTAGAAGCGCAGACCAGGCCTTTTGACCTCAAGAAGGATGTCTTCGTGCCTGATGACAAACAG
hum α ----------T-----G--C------C----C------A-TCGC-CT--GTG---------C--------GG-A

hum ß E  K  E  R  L  E  A  Q  T  R  P  F  D  L  K  K  D  V  F  V  P  D  D  K  Q
rab ß -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  E
hum α -  -  -  -  -  -  -  -  -  -  -  -  -  I  R  T  E  C  -  -  -  -  -  -  E
rab α -  -  -  -  -  -  -  -  -  -  -  -  -  I  R  T  E  R  -  -  -  -  -  -  E

                                                              end of exon ↓
hum ß GAGTTTGTCAAGGCCAAGATCGTGTCTCGAGAGGGTGGCAAAGTCACTGCCGAGACTGAGTATGGCAAG
hum α ----------A--------TT----C--G-----A-----G----T---T--A--C---A----G---

hum ß E  F  V  K  A  K  I  V  S  R  E  G  G  K  V  T  A  E  T  E  Y  G  K
rab ß -  -  -  -  -  -  -  M  -  -  -  -  -  -  -  -  -  -  -  -  H  -  -
hum α -  -  -  -  -  -  -  L  -  -  -  -  -  -  I  -  -  -  -  -  N  -  -
rab α -  -  -  E  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  N  -  -
```

Fig. 3. Sequence comparison of the first translated exon of the human β-MHC gene. Alignment of the human β-myosin HC-specific nucleotide and derived amino acid sequences of the first translated exon (exon 3) with corresponding sequences of the human α-, rabbit β-, and rabbit α-myosin HCs. *hum β*, Human β-myosinHC as derived from gMHC-3; *hum α*, human α-myosin HC, taken from Saez et al. (1987); *rab β* and *rab α*, amino acid sequence of rabbit β- and α-myosin HCs, derived from cloned DNA sequences (P. K. Umeda, personal communication). The *dashes* indicate nucleotides or amino acids matching with the human β-myosin HC. The *asterisks* in the sequence of α-myosin HC mark positions where a "best fit" alignment can only be achieved by the tentative introduction of nucleotides that were not reported in the original data (Saez et al. 1987). The generally highly conserved sequences of vertebrate myosin HCs suggest this addition (see Warrick and Spudich 1987). *CAGCCATGG*, translational start sequence (with one mismatch as compared with the reported consensus, see text) including the initiation codon ATG and the *Nco*I cleavage site

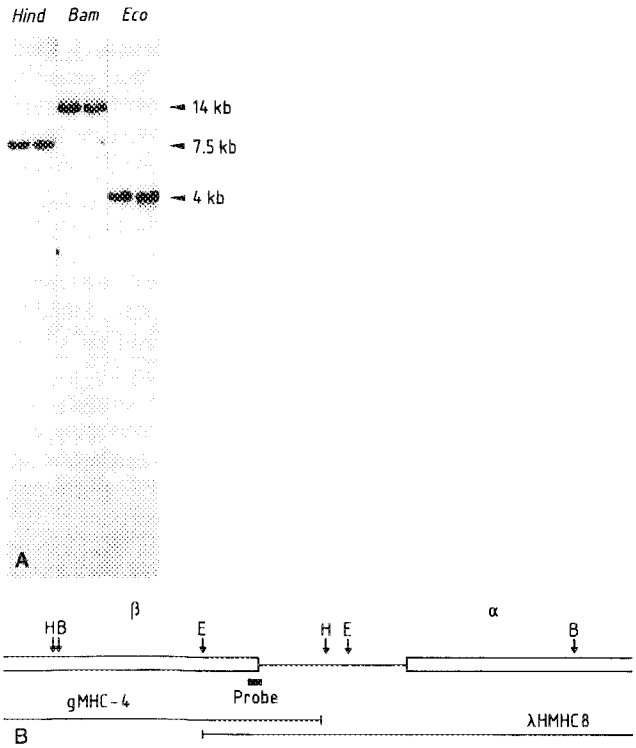


Fig. 4 A, B. Genomic restriction mapping with the β-myosin HC-specific probe EcoFIII. **A** Duplicated sets of 5 μg of human genomic DNA were restricted with *Hin*dIII, *Bam*HI, and *Eco*RI, separated on 1% agarose, blotted onto a nitrocellulose filter, and hybridized with ^{32}P-labeled EcoFIII DNA. The filter was washed at 68°C with 0.1 × SSC and exposed for 10 days. The lengths of the predominantly labeled bands are indicated. **B** The diagram depicts the restriction sites for *Hin*dIII *(H)*, *Bam*H1 *(B)*, and *Eco*RI *(R)* on the 3′ terminal half of clone gMHC-4 and on the 5′ terminal half of clone λHMHC8 (Saez et al. 1987). The origin of the probe is indicated

The 3′ terminus of the β-myosin HC gene has been identified by locating the AATAAA polyadenylation site close to the 3′ end of clone gMHC-1 (Lichter et al. 1986). That the entire coding sequence is included in this set of clones was shown by determining the translational start site on clones gMHC-2 and -3. This was achieved by sequencing a subregion that harbors a solitary *Nco*I cleavage site in a distance of about 22 kb from the 3′ end of the gene (Fig. 3). This site was found in identical positions on gMHC-2 and -3. In eukaryotic genes the translational initiation codon ATG relatively frequently forms part of a consensus sequence CCR*CCATG(G)* (Kozak 1984), which includes the recognition site for *Nco*I (underlined, see Fig. 3). Since the distance between the translation start and the poly(A) addition site in other known vertebrate myosin HC genes, i.e., the embryonic skeletal myosin HC genes of rat (Strehler et al. 1986) and chicken (Molina et al. 1987), are 22.4 and 20.4 kb long respectively, the *Nco*I site on gMHC-2 and -3 was a likely candidate to mark the beginning of the translated region.

Sequence data from this region of the human gene show that the initiation codon ATG and its immediate neighborhood conform to the above consensus sequence with the exception of the minus 4 position relative to ATG. We identified a putative 5′ splice site 8 bp 5′ to ATG. The downstream sequence included exon 3 with 67 amino acids of the N-terminus of the myosin heavy chain. (Exons 1 and 2 do not include coding information. Instead, they constitute, as will be shown elsewhere, the bulk of the 5′ nontranslated region of the β-myosin HC mRNA.)

It follows from a sequence comparison with other mammalian cardiac myosin HC sequences, both on DNA and protein level, that the protein sequence coded by this exon forms the amino terminus of β-myosin HC. The relevant sequences (Fig. 3) indicate close homology between the human and rabbit β-myosin HC sequences for the first coding exon: 94% for the protein and 92.8% for the DNA. (The unpublished rabbit data were kindly provided by P. K. Umeda, University of Chicago. The DNA sequence is not included in Fig. 3). Significantly less homologous is the human β-myosin HC to the corresponding amino terminus of either the human α-myosin HC (DNA, 75% protein, 77.6%; see Saez et al. 1987) or the rabbit α-myosin HC (81% on the protein level). The extent of protein sequence conservation between rabbit α- and β-myosin HC (83.6%) is not far from the degree of conservation between human α- and β-myosin HCs.

These data show positively that the coding information of the gene is completely contained within the cloned genomic region. Since this region extends about 5 kb beyond the ATG

initiation codon, we were able to identify the putative promotor by gene transfer experiments within a 1.3-kb *Hin*dIII fragment about 1.9 kb 5′ to the initiation start site (Horstmann-Herold et al. 1988). The promotor location is indicated in Fig. 1.

In addition, using EcoFIII DNA as a probe, we have located the restriction sites for *Hin*dIII, *Bam*H1 and *Eco*RI in genomic DNA within the 3′ end of the β-gene and adjacent downstream sequences. In blots (Southern 1975) we observed fragment lengths of 7.5, 14, and 4 kb, respectively (Fig. 4). These values fit into the combined restriction maps of clone gMHC-1 (or -4) and the recently reported human myosin HC clone λHMHC8 (Saez et al. 1987). These mapping data confirm that the β- and the α-myosin HC genes are closely linked in the genome.

Discussion

The β-type specificity of the myosin heavy chain gene, which we have cloned on a 30-kb-long genomic region, was originally determined by comparison of 3′ terminal translated and nontranslated sequences that are more similar to β-myosin heavy chain DNA and protein sequences than to other known mammalian myosin sequences (Lichter et al. 1986). We have extended this analysis to 5′ terminal sequences and find again a particularly close sequence relationship to nonhuman β-myosin sequences.

The β-myosin HC gene is expressed both in cardiac and in skeletal muscle. This conclusion rests on S1 nuclease protection mapping of mRNA with DNA probes from the 3′ end of this gene (Lichter et al. 1986; Saez et al. 1987; Jandreski et al. 1987) and also on evidence based on in vitro amplification of human myosin mRNA from skeletal and cardiac muscle in a polymerase chain reaction (Harbarth and Vosberg 1988). The β-myosin HC mRNA in human skeletal muscle is probably restricted to slow muscle fibers (Jandreski et al. 1987).

That the β-myosin HC gene is active in two different muscle tissues is corroborated by the independent isolation of identical β-myosin HC cDNA sequences from a cardiac as well as from a skeletal muscle cDNA library. We have isolated from a skeletal cDNA library a 4840-bp-long clone (K.W. Diederich, unpublished work) that carries at its 3′ end the nontranslated sequence that we have previously identified at the 3′ end of the genomic clone gMHC-1 (Lichter et al. 1986). We note that this part of the cDNA sequence is identical in its entire length with that obtained from others with a cardiac (Jandreski and Liew 1987) or skeletal muscle myosin HC cDNA (Saez and Leinwand 1986).

It has been suggested repeatedly (Bouvagnet et al. 1984; Jandreski et al. 1987; Mommaerts 1988; Tsuchimochi et al. 1988) that the β-myosin HC as a protein is not strictly unique, but exists in (at least) two different versions in mammalian cardiac and/or skeletal muscle. Such isoform variability of the β-myosin HC could theoretically be based on a closely related but nonallelic isogene, on alternative splicing of the transcript of one unique gene, or on posttranslational modification of the protein.

The β-specific probe EcoFIII, which carries the 3′ nontranslated sequence of β-myosin HC mRNA, hybridizes to only one fragment on genomic Southern blots after restriction with *Hin*dIII, *Bam*HI, and *Eco*RI (Fig. 4), indicating that the β-myosin HC gene is most probably a single-copy gene in the human genome. Thus, no evidence exists for a second, β-related gene. This is in agreement with the findings of Catanzaro and Morris (1986).

Alternative splicing of RNA could yield cDNAs with partially varying sequences. Differences in reported β-myosin HC cDNA sequences – pHMC3, from a cardiac source (Jandreski and Liew 1987), and pSMHCZ, from a skeletal source (Saez and Leinwand 1986) – have been taken to suggest that two differently spliced forms of β-myosin HC mRNA are present in human striated muscle (Jandreski et al. 1987). However, this notion is solely based on the comparison of independently derived DNA sequences from different laboratories. It was not confirmed by S1-mapping experiments with β-myosin HC-specific cDNA as a probe (Jandreski et al. 1987). The suggested alternative splicing would require putative exon sequences that were selectively incorporated into differentially spliced mRNA molecules. Exons able to appear variably in mRNA (e.g., in a tissue-specific manner) should exhibit a high degree of sequence homology. So far, our DNA sequence data covering about 80% of the entire β-myosin HC gene including the regions with the reported sequence differences do not provide evidence for exons which are potentially involved in alternative RNA processing (T. Jaenicke, K.W. Diederich, T. Ried, and H.-P. Vosberg, in preparation). Thus, the observed cDNA sequence differences may require another explanation.

With respect to covalent modifications of myosin heavy chains, at least one report indicates that posttranslational events (Bandman et al. 1982) may contribute in vivo to myosin HC isotype variability. The extent to which such mechanisms are used is, however, not known.

Different types of myosin HCs have in general rather similar structures (Warrick and Spudich 1987). We have found by comparing the amino termini of α- and β-myosin HCs from different mammalian species that within this frame of general similarities, specific subpatterns of sequence conservation exist among cardiac myosin HC genes. This observation corroborates the notion (Friedman et al. 1984; Lichter et al. 1986) that duplication of the cardiac myosin HC genes preceded the onset of mammalian species differentiation.

Genomic clones of the human cardiac myosin HC locus on chromosome 14 have been described by Appelhans and Vosberg (1983; clone designation, gMHC-1), by Catanzaro and Morris (1986; λHCMHC8, obviously resembling gMHC-1, and λHCMHC9, presumably part of the α-myosin HC gene), by Saez et al. (1987; λHMHC8, carrying N-terminal parts of the α-myosin HC gene in addition to the 3′ end of the β-myosin HC gene), and recently by Matsuoka et al. (1988; λgHMHCI, representing the C-terminal half of the α-myosin HC gene and – according to the respective restriction maps – resembling λHCMHC9). Our restriction mapping data obtained with genomic DNA (Fig. 4) show that the cloned 30-kb region represented by gMHC-1, -2, -3, and -4 overlaps with λHMHC8 (Saez et al. 1987). Since β-myosin HC exon sequences reported by these authors are identical with sequences that we have reported previously (Lichter et al. 1986), we conclude that a cloned contiguous region of at least 45 kb of the cardiac myosin HC locus is now available.

Acknowledgements. This work was supported by a grant from the Deutsche Forschungsgemeinschaft to K.W.D. We thank Dr. C. Sander for the computer program HD-MAXHOM and Dr. P.K. Umeda for communicating unpublished sequence information. We gratefully acknowledge the excellent technical assistance by Mrs. J. Schleich.

The restriction map for the clone λ gMHC-4 was determined by W. Haas.

References

Appelhans H, Vosberg HP (1983) Characterization of a human genomic DNA fragment coding for a myosin heavy chain. Hum Genet 65:198–203

Bandman E (1985) Myosin isozyme transitions in muscle development, maturation, and disease. Int Rev Cytol 97:97–131

Bandman E, Matsuda R, Strohman RC (1982) Myosin heavy chains from two different adult fast-twitch muscles have different peptide maps but identical mRNAs. Cell 29:645–650

Benton WD, Davis RW (1977) Screening λgt recombinant clones by hybridization to single plaques in situ. Science 196:180–182

Blau HM, Pavlath GK, Hardeman EC, Chiu CP, Silberstein L, Webster SG, Miller SC, Webster C (1985) Plasticity of the differentiated state. Science 230:758–766

Bouvagnet P, Leger J, Pons F, Dechesne C, Leger JJ (1984) Fiber types and myosin types in human atrial and ventricular myocardium: an anatomical description. Circ Res 55:794–804

Breitbart RE, Andreadis A, Nadal-Ginard B (1987) Alternative splicing: an ubiquitous mechanism for the generation of multiple protein isoforms from single genes. Annu Rev Biochem 56:467–495

Buckingham ME (1985) Actin and myosin multigene families: their expression during the formation of skeletal muscle. Essays Biochem 20:77–109

Butler-Browne GS, Whalen RG (1984) Myosin isozyme transitions occurring during the postnatal development of the rat soleus muscle. Dev Biol 102:324–334

Buttrick PM, Malhotra A, Brodman R, McDermott L, Lam L (1986) Myosin isoenzyme distribution in overloaded human atrial tissue. Circulation 74:477–483

Caplan AI, Fiszman MY, Eppenberger HM (1983) Molecular and cell isoforms during development. Science 221:921–927

Caruthers MH, Barone AD, Beaucage SL, Dodds DR, Fisher EF, McBride LJ, Matteucci M, Stabinsky Z, Tang J-Y (1987) Chemical synthesis of deoxyoligonucleotides by the phosphoramidite method. Methods Enzymol 154:287–313

Catanzaro DF, Morris BJ (1986) Human cardiac myosin heavy chain genes. Isolation of a genomic DNA clone and its characterization and of a second unique clone also present in the human genome. Circ Res 59:655–662

Clarke L, Carbon J (1976) A colony bank containing synthetic ColE1 hybrid plasmids representative of the entire *E. coli* genome. Cell 9:91–99

Ecob-Prince MS, Jenkison M, Butler-Browne GS, Whalen RG (1986) Neonatal and adult myosin heavy chain isoforms in a nerve-muscle culture system. J Cell Biol 103:995–1005

Edwards YH, Parker M, Povey S, West LF, Parrington MJ, Solomon E (1985) Human myosin heavy chain genes assigned to chromosome 17 using a human cDNA clone as a probe. Ann Hum Genet 49:101–109

Emerson CP, Bernstein SI (1987) Molecular genetics of myosin. Annu Rev Biochem 56:695–726

Everett AW, Umeda PK, Sinha AM, Rabinowitz M, Zak R (1986) Expression of myosin heavy chains during thyroid hormone-induced cardiac growth. Fed Proc 45:2568–2572

Feinberg AP, Vogelstein B (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132:6–13

Friedman DJ, Umeda PK, Sinha AM, Hsu HJ, Jakovcic S, Rabinowitz M (1984) Characterization of genomic clones specifying rabbit α- and β-ventricular myosin heavy chains. Proc Natl Acad Sci USA 81:3044–3048

Frischauf AM, Lehrach H, Poustka A, Murray N (1983) Lambda replacement vectors carrying polylinker sequences. J Mol Biol 170:827–842

Harbarth P, Vosberg HP (1988) Enzymatic amplification of myosin heavy chain mRNA sequences in vitro. DNA 7:297–306

Hiller G (1986) Lokalisierung des menschlichen beta-Myosingens auf Chromosom 14 und die Identifizierung von zwei mit diesem Gen gekoppelten Restriktionsfragmentlängenpolymorphismen. Thesis, University of Heidelberg

Hoffmann U, Axman C, Grisk A (1986) Myosin isoenzymes in normal and hypertrophied human hearts. Biomed Biochim Acta 45:985–996

Horstmann-Herold U, Lichter P, Mommaerts WFHM, Vosberg H-P (1988) Regulation of the human β-myosin heavy chain gene. In: Carraro U (ed) Sarcomeric and non-sarcomeric muscles: basic and applied research prospects for the 90s. Unipress, Padova (in press)

Izumo S, Nadal-Ginard B, Mahdavi V (1986) All members of the MHC multigene family respond to thyroid hormone in a highly tissue-specific manner. Science 231:597–600

Jandreski MA, Liew CC (1987) Construction of a human ventricular cDNA library and characterization of a beta myosin heavy chain cDNA clone. Hum Genet 76:47–53

Jandreski MA, Sole MJ, Liew CC (1987) Two different forms of beta myosin heavy chain are expressed in human striated muscle. Hum Genet 76:127–131

Karn J, Brenner S, Barnett L, Cesarini G (1980) Novel bacteriophage cloning vector. Proc Natl Acad Sci USA 77:5172–5176

Kozak M (1984) Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. Nucleic Acids Res 12:857–872

Leinwand LA, Saez L, McNally E, Nadal-Ginard B (1983a) Isolation and characterization of human myosin heavy chain genes. Proc Natl Acad Sci USA 80:3719–3720

Leinwand LA, Fournier REK, Nadal-Ginard B, Shows TB (1983b) Multigene family for sarcomeric myosin heavy chain in mouse and human DNA: localization on a single chromosome. Science 221:766–768

Lichter P (1986) Struktur und Expression genomischer Information für die schwere Kette eines menschlichen Myosins. Thesis, University of Heidelberg

Lichter P, Umeda PK, Levin JE, Vosberg HP (1986) Partial characterization of the human β-myosin heavy-chain gene which is expressed in heart and skeletal muscle. Eur J Biochem 160:419–426

Maniatis T, Fritsch EF, Sambrook J (eds) (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

Matsuoka R, Chambers A, Kimura M, Kanda N, Bruns G, Yoshida M, Takao A (1988) Molecular cloning and chromosomal localization of a gene for human cardiac myosin heavy-chain. Am J Med Genet 29:369–376

Maxam A, Gilbert W (1980) Sequencing end-labelled DNA with base specific chemical cleavage. Methods Enzymol 65:499–560

McDermott P, Daood M, Klein I (1985) Contraction regulates myosin synthesis and myosin content of cultured heart cells. Am J Physiol 249:H763–H769

Molina MI, Kropp KE, Gulick J, Robbins J (1987) The sequence of an embryonic myosin heavy chain gene and isolation of its corresponding cDNA. J Biol Chem 262:6478–6488

Mommaerts WFHM (1988) Fritz Lipman, a few personal memories, and: what else came out of the high-energy phosphate bond? In: Kleinkauf H, Döhren FR von, Jaenicke L (eds) The roots of modern biochemistry. de Gruyter, Berlin New York, pp 423–429

Nguyen HT, Gubits RM, Wydro RM, Nadal-Ginard B (1982) Sarcomeric myosin heavy chain is coded by a highly conserved multigene family. Proc Natl Acad Sci USA 79:5230–5234

Pette D, Vrbova G (1985) Neural control of phenotype expression in mammalian muscle fibers. Muscle Nerve 8:676–689

Rappold GA, Vosberg HP (1983) Chromosomal localization of a human myosin heavy chain gene by in situ hybridization. Hum Genet 65:195–197

Saez LJ, Leinwand LA (1986) Characterization of diverse forms of myosin heavy chain expressed in adult human skeletal muscle. Nucleic Acids Res 14:2951–2969

Saez LJ, Gianola KM, McNally EM, Feghali R, Eddy R, Shows TB, Leinwand LA (1987) Human cardiac myosin heavy chain genes and their linkage in the genome. Nucleic Acids Res 15:5443–5459

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain termination inhibitors. Proc Natl Acad Sci USA 74:5463–5467

Silhavy TJ, Berman ML, Enquist LW (1984) Experiments with gene fusions. Cold Spring Habor Laboratory, Cold Spring Habor, NY

Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98:503–517

Strehler EE, Strehler-Page MA, Perriard JC, Periasamy M, Nadal-Ginard B (1986) Complete nucleotide and encoded amino acid sequence of a mammalian myosin heavy chain gene – evidence against intron-dependent evolution of the rod. J Mol Biol 190: 291–317

Tsuchimochi H, Kuro-o M, Koyama H, Kurabayashi M, Sugi M, Takaku F, Furuta S, Yazaki Y (1988) Heterogeneity of β-type myosin isozymes in the human heart and regulational mechanisms in their expression. Immunohistochemical study using monoclonal antibodies. J Clin Invest 81:110–118

Warrick HM, Spudich JA (1987) Myosin structure and function in cell motility. Annu Rev Cell Biol 3:379–421

Weydert A, Daubas P, Caravatti M, Minty A, Bugaisky G, Cohen A, Robert B, Buckingham ME (1983) Sequential accumulation of mRNAs encoding different myosin heavy chain isoforms during skeletal muscle development in vivo detected with a recombinant plasmid identified as coding for an adult fast myosin heavy chain from mouse skeletal muscle. J Biol Chem 258:13867–13874

Whalen RG, Toutant M, Butler-Browne GS, Watkins SC (1985) Hereditary pituitary dwarfism in mice affects skeletal and cardiac myosin isozyme transitions differently. J Cell Biol 101:603–609

Received June 7, 1988 / Revised August 21, 1988